



**PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE  
FACULTAD DE CIENCIAS SOCIALES**

**Escuela de Psicología**

**Curso: Métodos de Investigación Cuantitativa (PSI2301)**

**Sección: 1**

**Docente: Diego Carrasco**

**Ayudantes: Carolina Castillo, Catalina Miranda, y Elisa Rodríguez,**

### **Instrucciones Tarea 5**

Las tareas individuales consisten en realizar una serie de análisis estadísticos simples en R y entregar un pequeño informe con los resultados en formato Word o en PDF (ideal no más de una plana), incluyendo su nombre completo y el número de tarea al que corresponde el archivo.

Adicionalmente, deberán entregar el archivo de R con el código que se utilizó para realizar los análisis de la tarea en el buzón de Canvas que se creará para dicho propósito. Los resultados que usted presente serán verificados con los resultados del código que entregue.

Tanto el archivo Word como el script de R que escribieron para realizar la tarea deben ser entregados a más tardar el **Miércoles 22 de junio de 2022 a las 23:59** en el buzón de Canvas que será creado para dicho propósito.

#### **Base de datos**

La base de datos con la que trabajara proviene del estudio International Civic and Citizenship Education Study de 2016<sup>1</sup>, el cual evalúa el rendimiento en educación cívica y ciudadana de los estudiantes de 24 países. Esta base de datos posee una submuestra de datos representativos de estudiantes de octavo grado de Chile, Colombia, República Dominicana, México, y Perú.

Las variables originales de la base de datos pública han sido cambiadas, de modo que sus nombres sean más cortos. Para revisar más detalles de las variables contenidas en la base de datos, ver:  
[https://github.com/dacarras/psi2301\\_examples/blob/master/data/civ\\_data\\_codebook.pdf](https://github.com/dacarras/psi2301_examples/blob/master/data/civ_data_codebook.pdf)

En la Tabla 1 se presenta el listado de variables contenidas en la base de datos que usted tendrá que usar para desarrollar la tarea:

---

<sup>1</sup> Para más información ver página del estudio: <https://www.iea.nl/studies/iea/iccs/2016>

**Tabla 1 -***Listado de variables disponibles en la base de datos*

Variable	Clase	Descripción	Valores
ge1	numérica	Los hombres y las mujeres deberían tener las mismas oportunidades de participar en el gobierno.	Para todos los ítems ge1 a ge6:  1=Muy de acuerdo, 2=De acuerdo, 3=En desacuerdo, 4=Muy en desacuerdo
ge2	numérica	Los hombres y las mujeres deberían tener los mismos derechos en todo sentido.	
ge3	numérica	Los hombres y las mujeres deberían obtener el mismo pago si es que realizan el mismo trabajo.	
ge4	numérica	Las mujeres no deberían participar de la política.	
ge5	numérica	Cuando no hay muchos trabajos disponibles, los hombres deberían tener más derecho a trabajar que las mujeres.	
ge6	numérica	Los hombres están mejor calificados para ser líderes políticos que las mujeres	
ses	numérica	Escala de Nivel Socioeconómico de las familias	Escala estandarizada a Z, con media 0, desviación estándar 1, donde mayor valor mayor nivel socioeconómico.
edu	cualitativa	Nivel de Educación máximo alcanzado por los padres	“tertiary” = Educación terciaria (universitaria, y otros grados de educación superior; “non-tertiary” = Sin educación superior
civ_level	cualitativa	Nivel de Conocimiento Cívico	“level 3” = Alta sofisticación, “level 2” = sofisticación promedio, “level 1” = sofisticación baja; “level 0” = sofisticación menos que baja
cty	cualitativa	País	“CHL” = Chile, “COL” = Colombia, “DOM” = República Dominicana, “MEX” = México, “PER” = Perú

## Selección de la submuestra de la base de datos con que deberá realizar la tarea

Para esta tarea seleccionaremos una submuestra aleatoria estratificada por país (cty), para lo cual deberá comenzar el archivo de código con el siguiente conjunto de comandos teniendo cuidado de reemplazar el valor 12345678 con su propio RUT sin dígito verificador:

```
# abrir datos originales
civ_data <- read.csv(
url('https://raw.githubusercontent.com/dacarras/psi2301_examples/master/data/civ_data.csv')
)

# crear datos para usuario
# fijar semilla, para crear datos
set.seed(123456) # reemplazar este valor con su rut
library(dplyr)

data_tarea_5 <- civ_data %>%
  group_by(cty) %>%
  sample_n(500 , replace = FALSE) %>%
  ungroup() %>%
  dplyr::glimpse()

# Nota: Seleccionaremos el 500 casos por país

# Verificamos que nuestra nueva submuestra tiene 2500 casos

nrow(data_tarea_5)
# ¡Recuerde que debe usar los datos en el objeto data_tarea_5 para hacer la tarea!
```

## Ejercicios

### Sección 1

En las siguientes preguntas, se le pedirá responder la siguiente pregunta ¿Cuál es la asociación entre el nivel de Conocimiento Cívico de los estudiantes (civ\_level) y el máximo nivel de educación alcanzado por sus padres” (edu)? Para dar cuenta de esta pregunta, usted debe evaluar la asociación entre el “Nivel de Conocimiento Cívico” (civ\_level) de los estudiantes y el “Nivel de Educación máximo alcanzado por los padres” (edu). El problema está dividido en una secuencia de preguntas que usted debe resolver.

1.1 Construya y presente una tabla de contingencia que muestre los porcentajes de la variable “Nivel de Conocimiento Cívico” (civ\_level) en las filas y los porcentajes de la variable “Nivel de Educación máximo alcanzado por los padres” (edu) en las columnas (1 pto.).

1.2 ¿Qué prueba estadística emplearía para evaluar esta asociación? ¿Por qué? (1 pto.)

1.3 Calcule la prueba estadística que indicó en 1.2, y describa en palabras el resultado obtenido reportando el estadístico, con sus grados de libertad y la significancia estadística. En función de la evidencia encontrada en la prueba estadística y en la tabla de contingencia construida en 1.1, interprete sustantivamente los resultados y concluya acerca de la posible asociación entre el nivel de Conocimiento Cívico de los estudiantes (civ\_level) y el máximo nivel de educación alcanzado por sus padres (2 ptos.).

## Sección 2

En la siguiente sección de esta tarea, se le pedirá que evalúe la confiabilidad de un conjunto de respuestas a ítems referidos al apoyo a la igualdad de derechos entre hombres y mujeres (ge1-ge6), también llamada “Actitudes de los estudiantes a la igualdad de género”. Para evaluar la confiabilidad que presentan estas respuestas en un puntaje de medias de respuestas usted tendrá que abordar una serie de preguntas.

2.1 Si a usted le encargaran evaluar la confiabilidad de una escala creada a partir de las respuestas a los ítems ge1-ge6. ¿Qué procedimiento estadístico emplearía para evaluarla? ¿Por qué? (1 pto.)

2.2 ¿Existen en esta escala ítems invertidos? Identifíquelos. En esta clasificación, considere que la dirección deseada de los puntajes es que mayor puntaje representen mayor apoyo a la igualdad de derechos de hombres y mujeres. (0.5 pto.)

2.3 Estime la confiabilidad de la escala generada con las respuestas a los ítems ge1-ge6, empleando el análisis que usted indicó en 2.1. Reporte el resultado obtenido e indique el grado confiabilidad alcanzado (2 ptos.).

2.4 De acuerdo con la evidencia obtenida en el análisis, ¿Es posible aumentar la confiabilidad de este instrumento, eliminando uno o más ítems de la escala “Actitudes de los estudiantes a la igualdad de género”? Justifique su respuesta (1 pto.).

2.5 Calcule un puntaje promedio con los ítems ge1-ge6, de modo tal que mayor valor represente mayor apoyo a la igualdad de derechos de hombres y mujeres; y luego calcule una correlación de este puntaje creado con el nivel socioeconómico de las familias de los estudiantes. Describa los resultados obtenidos. En su respuesta, indique el coeficiente de correlación obtenida, y describa los resultados de forma sustantiva (3 ptos.) Nota: recuerde invertir los puntajes de los ítems respectivos.